MoodExplorer: Towards Compound Emotion Detection via Smartphone Sensing

XIAO ZHANG, Nanjing University, China
WENZHONG LI*, Nanjing University, China
XU CHEN, Sun Yat-sen University, China
SANGLU LU, Nanjing University, China

Social psychology and neuroscience had confirmed that emotion state exerts a significant effect on human communication, perception, social behavior and decision making. With the wide availability of smartphones equipped with microphone, accelerometer, GPS, and other source of sensors, it is worthwhile to explore the possibility of automatic emotion detection via smartphone sensing. Particularly, we focus on a novel research problem that tries to detect the compound emotion (a set of multiple dimensional basic emotions) of smartphone users. We observe that users' self-reported emotional states have high correlation with their smartphone usage patterns and sensing data. Based on the observations, we exploit a feature extraction and selection algorithm to find the most significant features. We further adopt a factor graph model to tackle the correlations between features and emotion labels, and propose a machine learning algorithm for compound emotion detection based on the smartphone sensing data. The proposed mechanism is implemented as an APP called MoodExplorer in Android platform. Extensive experiments conducted on the smartphone data collected from 30 university students show that MoodExplorer can recognize users' compound emotions with 76.0% exact match on average.

CCS Concepts: • **Human-centered computing** → **Smartphones**; *Ubiquitous and mobile computing systems and tools*;

Additional Key Words and Phrases: Emotion detection, Compound emotion, Smartphone sensing, Factor graph

ACM Reference Format:
Xiao Zhang, Wenzhong Li, Xu Chen, and Sanglu Lu. 2017. MoodExplorer: Towards Compound Emotion Detection via Smartphone Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 176 (December 2017), 30 pages. https://doi.org/10.1145/3161414

1 INTRODUCTION

Mood sensing is receiving widespread attention from the social psychology, neuroscience, and computer science in the past years [14][30][52]. The mood or emotion state plays an important role in human daily lives, which has great influence on people's communication, perception, social behavior, and decision

*The corresponding author is Wenzhong Li, Email lwz@nju.edu.cn.

Authors' addresses: Xiao Zhang, Nanjing University, State Key Laboratory for Novel Software Technology, Nanjing, Jiangsu, 210023, China; Wenzhong Li, Nanjing University, State Key Laboratory for Novel Software Technology, Nanjing, Jiangsu, 210023, China; Xu Chen, Sun Yat-sen University, School of Data and Computer Science, Guangzhou, Guangdong, China; Sanglu Lu, Nanjing University, State Key Laboratory for Novel Software Technology, Nanjing, Jiangsu, 210023, China.

2474-9567/2017/12-ART176 $15.00
https://doi.org/10.1145/3161414

Fig. 1. Emotion detection via smartphone sensing.

making. Automatic mood detection is a challenging task, which envisions a wide range of new mood-aware application scenarios. For instance, people enjoy different styles of music and movie, which not only depends on their preference, but also relates to their mood and personality. Therefore a smart recommendation system should take into account the mood to enhance user experience. Another example is advertisement, which is found interesting or annoying for different persons with different emotion states. So advertisement can be more effective and personalized if peoples' feelings can be considered. With the prevalence of social network applications, sharing the mood among family and close friends can help people to strengthen their bond and improve the way of social communication. Furthermore, robots will be widely used in the near future in different aspects of our lives. The robots can be more intelligent and humanized if they can "read the mood" of the human they work for. Last but not least, understanding the emotion state and its evolution is important to evaluate individual's psychological health and mental well-being [10][19].

Mood detection had been reported using body physiological signal such as heart beat, blood pressure, breath rate, etc [17][26][53]. However, monitoring such signals relies on expensive dedicated devices, which is infeasible for pervasive device-free detection. Some existing works sought to recognize emotion by audio and video signals [49]. For example, Ang et al. explored speech-based recognition for the emotion of annoyance and frustration [1]. Ashraf et al. detected pain expression by recognition of facial signals [3]. The MoodMeter proposed the recognition of smiling face via campus video cameras [22]. However, visual feature recognition only reflects people's expression in a snapshot. Besides, collecting and analyzing audio and video data is a high-computational task, and such signals cannot be captured everywhere without wide deployment of cameras.

Nowadays smartphones are widely used in people's daily lives for business, social and entertainment purposes [4][30]. As shown in Fig. 1, there are many sensors embedded in modern smartphones: microphone, accelerometer, electronic compass, GPS, proximity, etc. The data captured from the sensors possesses profound information and can be exploited to infer user's social behaviors such as physical movement, social communications, location, etc. Intuitively, the usage of smartphones and the context information is correlated to users' emotions. For example, people may play games with cellphone when they are happy, or they may feel depressed if they live in a noisy environment. Motivated by this, several works employed smartphone sensing data for emotion detection. Bogomolov et al. proposed a multifactorial statistical model to recognize daily stress by comprehensively analyzing mobile phone data and weather conditions [6]. Sun et al. employed sensor data, APP usage information, and SMS content for cold-start emotion prediction using transfer learning [43]. However, their approach is content-based, which requires users' highly privacy-sensitive information such as SMS content. LiKamWa et al. proposed a system for mood detection utilizing Email, SMS, location, and App usage duration as features [30]. What needs to be stressed is that the literature only considered single emotion detection, which assumes that people is in only one emotion state in a period. Different from the existing works, we weaken such assumption and study the co-existence of multiple emotions called compound emotion. According to Plutchik's theory [34], emotion is not necessary in a pure state and could be the mixture of basic emotions. The study of [11][55] also showed compound facial emotions in human facial expressions. For instance, "happily surprised" is a compound emotional expression that combines basic emotions of happiness and surprise. Different from the compound facial emotions that individuals should experience at the same time, the compound emotion studied in our paper refers to a set of basic emotions that an individual experienced in a duration, which can occur simultaneously or alternately. In this paper, we provide formal expression of compound emotion as a vector of multiple basic emotions with discrete levels, and propose a machine learning algorithm to detect compound emotion. *To the best of our knowledge, this is the first work of compound emotion detection based on smartphone sensing, which has not been addressed in the past.*

Specially, we propose a system called *MoodExplorer* to enable compound emotion detection via cellphone. We first develop an Android APP to allow people to report their emotion states and collect the data of smartphone sensors and usage patterns. The APP was installed and tested by 30 university students for a month, which forms the dataset for model training and testing. It is observed in the dataset that about 60% reported emotions are compound emotions consisting of more than two basic emotions, which verifies the motivation of our work. Based on the dataset, we extract different types of features in respect of *environment*, *contact*, *APP usage*, and *human activities*, which are used for automatic emotion detection. To best tune the performance, we present a feature selection algorithm using the *problem-transformation approach* and *ReliefF measure* to choose the most significant features. Using the selected features as input, we adopt the *factor graph model* to represent the correlations between features and multiple emotion labels, and propose a learning algorithm for compound emotion detection. We conduct extensive experiments on the collected dataset, which demonstrates that the exact match of compound emotion achieves 76.0% on average.

The main contributions of the paper are summarized as follows.

- We identify the compound emotion detection problem. Unlike existing works that assume human emotion is exclusive, we observe that people usually report their emotion states as the combination of several basic emotions. The compound emotion is found highly correlated with users' smartphone usage patterns and sensing data, therefore compound emotion detection is possible without the need of knowing people's body physiological signals or facial expressions. To the best of our knowledge, smartphone-based compound emotion detection has not been addressed in the literature.
- We propose a novel compound emotion detection method using smartphone sensing. Specifically, we extract different types of features from the sensing data to show the environment, social contact, APP usage and activities of individuals, and apply feature selection algorithm to find the most significant features. Using the selected features as input, we propose a machine learning algorithm to derive the probable emotion labels by maximizing a posteriori probability. The proposed machine

learning algorithm uses a factor graph to depict the correlations between features and emotion labels and the correlations across different basic emotions, which is shown to be suitable for compound emotion detection.

- We develop and implement the MoodExplorer system for automatic emotion detection. The proposed system is implemented as an Android APP, which was tested by 30 university students. Based on the sensing data collected by the smartphones and the emotion states reported by the participants, we train the machine learning model and test its performance. It is shown that compound emotion can be correctly detected by our system with 76.0% exact match on average.

The rest of the paper is organized as follows. Section 2 provides an introduction to the related work about affect measurement models and emotion recognition mechanisms. Section 3 proposes the compound emotion model and formalizes the compound emotion detection problem. Section 4 introduces the system design and data collection process. Section 5 proposes data processing method including feature extraction and feature selection, and conducts correlation analysis of the selected features. Section 6 develops an efficient factor graph model for compound emotion detection. Section 7 reports experimental results, and demonstrates the performance of the proposed factor graph method. Section 8 concludes the work.

2 RELATED WORK

2.1 Models for emotion and mood measurement

Emotion and mood have been widely studied in psychology, sociology, and neuroscience [35]. Generally, "emotion" refers to the current instantaneous feeling, and "mood" refers to the average feelings over a longer period of time. A variety of models have been proposed to measure and quantify emotion and mood. Such models can be applied to measure both instantaneous and long-term feelings, and they can be used for emotion and mood measurement without restrict distinction.

One frequently used measure for general affective states is the Positive and Negative Affect Schedule (PANAS) model [9]. Participants completing the PANAS are asked to rate the extent to which they experienced each out of 20 emotions on a 5-point Likert Scale ranging from "very slightly" to "very much". Participants may be asked how they feel right now or during longer periods of time (e.g. during the past month) according to the purpose of the measurement of emotion or mood.

The discrete category model [12][45] described emotion through a set of categories. One of the most popular model is Ekman's six basic categories [12]: happiness, sadness, anger, surprise, fear, disgust. The Ekman's model is intuitive and understandable to normal users, and it allows the co-existence of multiple emotions with different intensive levels. With its high applicability, the Ekman's emotion model was widely adopted by many studies [36] [55].

The Circumplex mood model [35] employed a two-dimensional circumplex to represent the emotional state of the participants: the pleasure dimension measures the degree of positive and negative feelings, and the activeness dimension measures the likelihood for a user to take action under the mood state. Each dimension is quantified using a score, hence the Circumplex mood model provides continuous measurement of the mood, and has been adopted in many studies [30][43].

2.2 Emotion recognition

Emotion recognition is a fine-grained sentiment analysis, which aims to identify emotions from various information resources, such as video, image, text and so on. In the recent years, a large number of works have been done on this area. In the work of [22], the researchers identified smile faces through the camera in the campus. The study found that the users' emotional states had a cyclical pattern and were highly correlated with external events. In the traditional emotion classification, the emotional state is generally classified into one discrete category. Some works adopted the continuous probability distribution to depict the emotional state of an image, and proposed to use the Gaussian mixture model [54] for emotion recognition. As for emotion analysis based on text [28], the authors classified sentence-level emotion considering label context dependence, and formalized the problem as a multi-label classification problem, which allowed the detection of several emotions in a single sentence. Recently, the EQ-Radio proposed the usage of wireless signals to monitor individual's heartbeats, which was further used as features for emotion detection [53]. However, their work relied on dedicate WiFi device and on-body sensors such as ECG monitors.

2.3 Detection of human mental well-being based on smartphones

Nowadays, with the rapid adoption of smartphones, researchers have shown that it is possible to adopt smartphone sensing data to infer and detect human mental well-being such as stress, anxiety, depression, emotion, and mood [8][31][37][41][42][51].

DeepMood [8] detected bipolar affective disorder utilizing typing dynamics and accelerometer sensor data in the smartphone based on multi-view neural network. Herdem et al. [21] aimed to help mobile individuals to interact offline with friends when they need emotional support. Bogomolov et al. proposed a multifactorial statistical model to recognize daily stress by comprehensively analyzing mobile phone data and weather conditions [6]. Canzian et al. monitored the depression states of users by means of smartphone mobility trace analysis [7]. Mottelson et al. proposed the detection of positive and negative affect using mobile commodity sensors in the wild [32]. The iSelf system [43] employed sensor data, APP usage information, and SMS content for emotion prediction using transfer learning. MoodScope [30] utilized Email, SMS contact information, website visiting information, location, and App usage duration to detect users' mood. However, the existing works only consider single emotion detection, which ignored the fact that multiple emotions may co-exist in a period. In our paper, we make the first attempt to solve the compound emotion detection problem in smartphone sensing environment, which corresponds to the derivation of a multi-dimensional emotion vector with discrete levels leveraging a set of smartphone-generated sensing information.

3 COMPOUND EMOTION MODEL

3.1 Definitions

The phrases *mood* and *emotion* are highly related but have slight difference in several aspects [5]. Generally speaking, emotion is instantaneous intensive feeling reacting to some events, while mood can be viewed as long-lasting internal emotion state of an individual [16]. Since the state-of-the-art smartphone can not capture every human instantaneous activities and events so far, detecting instantaneous emotions is unlikely and impracticable. Therefore our study is interested in the measurement of short-term mood in a time interval of several hours. Unlike the conventional definition of mood, we particularly focus on the detection of *compound emotion*, which is defined as follows.

Definition 3.1 (Compound emotion). The compound emotion is a set of emotions that an individual experienced in a short duration. The duration is measured by a time interval typically within a few hours, during which the multiple emotions can occur simultaneously or successively.

According to the definition, compound emotion is different from the conventional concept of emotion since it allows the co-existence of multiple emotions. Compound emotion is also different from the conventional concept of mood since it less emphasizes on the long-lasting emotion states but more focused on the measurement of affect in a short duration.

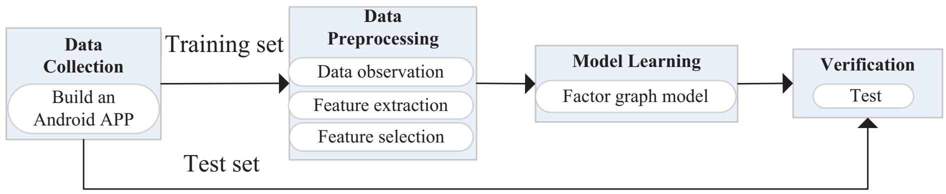


Fig. 2. MoodExplorer system framework.

The “compound emotion” is not a standard term in the literature, so we introduce our own definition in the paper. Some previous works also showed the co-existence of multiple emotions in human facial expression [11]. For example, the study of [11] showed 21 combined facial emotions including “happily surprised”, “angrily surprised”, “sadly feared”, “sadly disgusted”, etc, which are identified from the human facial expressions according to their images. To avoid confusion, we refer to their findings as “compound facial emotions”. The compound facial emotion describes multiple emotions occurring at the same time, while our definition of compound emotion measures multiple emotions in a duration, which can occur simultaneously or at different times.

The detection of compound emotion using smartphone sensing is based on an implicit assumption that individuals’ feelings are correlated to their smartphone usage. Such correlation is rather intuitive. For example, one may feel happy when she shops online, or one may call a close friend when she feels sad. Although the cause-effect of emotion-smartphone correlation is hard to depict, we can exploit the correlation to infer individual’s compound emotion by exploring the smartphone sensing data. The correlation of the compound emotion and the smartphone sensing data will be discussed in Section 5.

The widespread Ekman’s discrete category model [12] consists of six basic emotion categories: happiness, sadness, anger, surprise, fear, and disgust. In this paper, we adopt the idea of the combination of Ekman’s six basic emotion categories to express compound emotions. Specifically, we assign a discrete score for each basic emotion category, and represent the *compound emotion state* of an individual by a 6-tuple vector

$$Y =< S_{happy}, S_{sad}, S_{anger}, S_{surprise}, S_{fear}, S_{disgust} >, \tag{1}$$

where S_* is the score of the corresponding basic emotion category, which is quantified to 5 levels in $\{1, 2, 3, 4, 5\}$ representing *null*, *slight*, *moderate*, *strong*, and *extreme* respectively.

3.2 Compound emotion detection problem

We address the problem of compound emotion detection based on the sensing data collected from the smartphone. Given the massive source of sensing information from various sensors including microphone, accelerometer, electronic compass, light sensor, etc., as well as the logged APP usage history, it is a challenging task to infer compound emotion with the multimodality data.

To achieve efficient compound emotion detection, we first need to extract the most useful information known as *features* from the raw data. Given the obtained vector of features denoted by X, the task is to identify the user’s compound emotion. Particulary, we want to find a function to map the feature vector X to the compound emotion presentation Y:

$$f : (X) \to Y, \tag{2}$$

where Y is a multi-dimensional emotion vector as defined in Eq. (1).

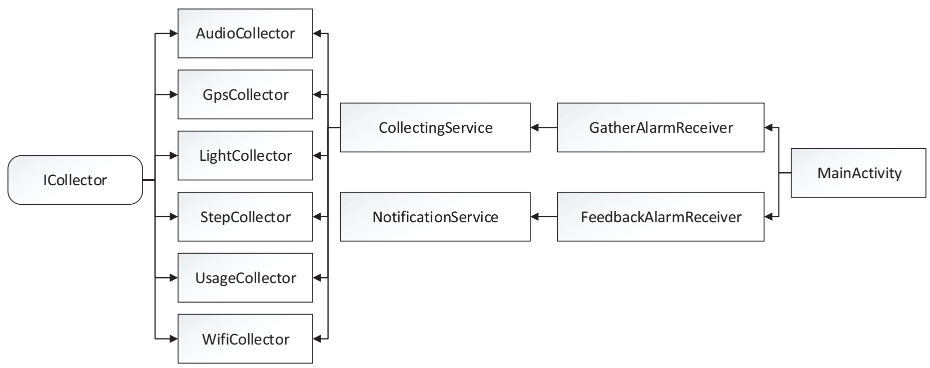


Fig. 3. Class structure of the MoodExplorer APP.

In the following sections, we will propose solution framework and methodology to solve the compound emotion detection problem.

4 SYSTEM DESIGN AND DATA COLLECTION

4.1 System framework

We propose the system framework as shown in Fig. 2 for compound emotion detection via smartphone sensing. First, we develop an Android APP to collect smartphone data from the users. The APP allows users to report their compound emotions to the server periodically, which form the ground truth and are used to build training set and test set. Then we extract features from the collected data. Since the number of the extracted features is large, we further apply a feature selection method to choose the most significant features to form the feature vector. Using the feature vectors and the labeled instances as input, we train a compound emotion detection model based on factor graph. Finally, we use the test set with users' reported emotions to verify the performance of the proposed system.

The detailed techniques are introduced in the following subsections.

4.2 Implementation of MoodExplorer

We implement the APP on Android platform for collecting sensing data and compound emotion reports. Fig. 4 shows the screen shots of the MoodExplorer APP. It enables several functions such as reporting the compound emotions, searching the history of reports, and showing the statistics of the cellphone sensing data and APP usages. The APP sends three notifications per day, which are at 10AM, 3PM, and 8PM respectively, at least 5 hours apart from one another. Users can report their emotional states experienced in the past few hours either launching the notification or opening the APP directly. The APP adopts the Ekman's six class basic emotion model and asks users to evaluate the intensity level for each of the basic emotion category in the scale of 1 to 5.

In detail, the class structure of the APP is shown in Fig. 3. The data collection process *ICollector* runs in the background and it is invoked by the system *AlarmManager* periodically. To save energy, the data collection interval is set to 5 minutes. Each time the ICollector is invoked, it will read the smartphone sensors. Particularly, it reads the GPS location of the smartphone; checks the on/off state of the smartphone screen; scans the WiFi signals nearby to log the IDs of scanned access points (APs)

Fig. 4. Screenshots of MoodExplorer.

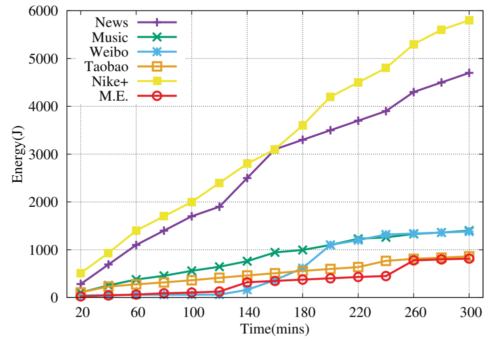


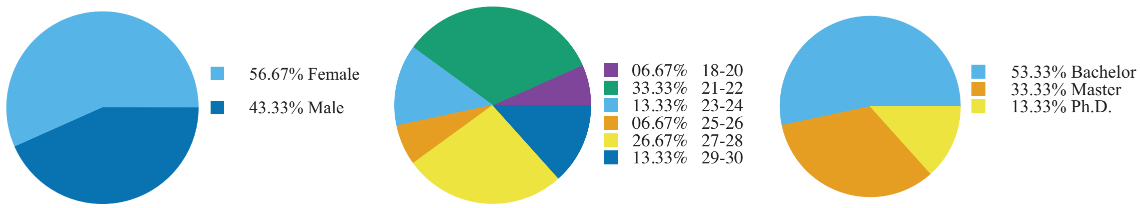
Fig. 5. Energy consumption of different Apps.

and the receive signal strength (RSS); and then it reads the microphone, the light sensor and the accelerometer, electronic compass, gyroscope measurements for 15 seconds and records the data in a local database. The users' APP usage information and social activities, which include opening/closing an APP, making/answering a phone call, sending/receiving a SMS, are also logged as events with timestamps in the system. The recorded data is stored in the smartphone's file system and submitted to a remote cloud server daily for analysis. To save the communication cost, the data is uploaded only when there is available WiFi connection.

The APP does not record privacy-sensitive information such as the content of the SMS and the voice during a call. All cellphone numbers and user names are anonymized by mapping them to random IDs. The SSIDs of the APs are also mapped to random IDs to conceal their real name. The GPS coordinates of the users are collected under the users' permission (the APP will ask for the authorization to use the GPS data when the APP is installed). If a user concerns about her location privacy, she can simply turn-off the GPS or unauthorise GPS to MoodExplorer to avoid being tracked by our APP. For the APP usage, we only record the category (e.g., shopping, entertainment, social networks, etc) of the APPs without recording their actual names. For the sensor data such as microphone sound and light, we compute their mean, variance, and other measurements in a duration as introduced in the feature extraction, and submit the feature measurements to the server without keeping the original data.

4.3 Energy consumption of MoodExplorer

Energy consumption is one important issue for smartphone. To evaluate the energy consumption, we compare the energy consumption of MoodExplorer with several other widely used applications: a news APP (Today' Tops), a music APP (Kugou Music), a social network APP (Weibo), a shopping APP (Taobao), and a fitness APP (Nike+). We run the APPs for 5 hours and compare their energy consumption according to the battery logs. The results are shown in Fig. 5. As shown in the figure, the energy consumption of MoodExplorer is not severe, which is very close to the shopping APP (Taobao) and the social network APP (Weibo). It consumes only half energy of the music APP (Kugou Music). The news APP (Today' Tops) and fitness APP (Nike+) consumes much more energy than MoodExplorer since they may acquire sensor data and network communications frequently. The experiments show that sampling sensor data every 5 minutes in MoodExplorer does not affect the usage of smartphones and the energy consumption is in an acceptable level.



(a) Gender. (b) Age. (c) Education.

Fig. 6. Demographics of the 30 participants.

4.4 Data collection

To test the MoodExplorer APP, we recruit 42 student volunteers to participate in the experiment of data collection. The students were asked to install the MoodExplorer APP in their smartphones, and to utilize the APP to report their emotions three times per day around 10AM, 3PM, and 8PM. We did not restrict the time to report their emotions, but we required that the interval of consecutive reports should be longer than 5 hours. The student volunteers were studying computer science, engineering, and business in our university, who had adequate computer skills and utilized smartphones actively. They were told about the research purpose of the experiments and were aware of that their data will be used for emotion study. To encourage the students to involve in the experiments, we sent them a thank-you gift (e.g., a USB flash drive, a mobile MicroSD memory card, a T-shirt, etc) after they submitted enough number of emotion reports. The data collection lasted for about one month, from June. 15 to July. 14, 2016. According to the returned results, we have 30 students that submitted emotion reports more than 50 times, which are used in our analysis.

Among the 30 students, there are about 57% females and 43% males. Their ages are from 18 to 30, and are mostly concentrated in the range of 21-24 and 27-30. There are about 53% undergraduates, 33% master students, and 14% PhD students. The demographics of the participants are shown in Fig. 6.

5 FEATURE EXTRACTION AND SELECTION

5.1 Data observation

After data collection, we need to preprocess the dataset. We mainly remove the unqualified submissions which caused by missing sensing data or compound emotion labels. If a user failed to submit the report in an interval, then the corresponding data will not be used for the model due to the lack of label. If part of the sensing data is missing in an instance, for example, a user may disable the GPS in some time intervals, then the corresponding feature will be set to "null". The null features will not affect the model and detection accuracy much since they are unlikely to be chosen to build the model after feature selection.

We make some observations to the collected dataset after data preprocessing. We first observe the distribution of number of qualified submissions, which is shown in Fig. 7. As shown in the figure, there are about 57% (17) users submitting qualified compound emotion reports more than 80 times, and 2 users continuing to submit reports after one month. Most of the participants submitted more than 60 qualified reports, which are used to form the training sets and test sets for our model.

Then we study the number of emotion labels reported by the individuals, which distribution is shown in Fig. 8. As shown in the figure, about 41% reports have single kind of emotion; about 40% reports contain two kinds of emotions, and about 13% contain 3 kinds of emotions. In another words, there are near 60% reported instances that are compound (multi-dimensional) emotions. This illustrates that compound emotions are very common in daily life, which verifies the key motivation of our study.

Removing the "null" emotions, we also show the distribution of different level of the reported compound emotions in Fig. 10. As shown in the figure, happiness is the most common emotion that contains in the user reports, where about 14% are moderate, about 11% are slight, about 7% are strong and about 4% are extreme. Sadness is the second common emotion, which yields about 10.5% are slight and about 5% are moderate. Other emotions such as anger, surprise, fear and disgust are less common, and most of them are less than 10% in the total user reports.

We further investigate the combination of basic emotions. If two basic emotions co-exist in the same report, we consider them correlated. We calculate the fraction of correlated emotions in the users' reports, which is illustrated in Fig. 9, where thicker edge corresponds to higher frequency of co-existence. As shown in the figure, several emotion pairs, such as *sadness&disgust*, *sadness&fear*, *happiness&surprise*, frequently appear together. This suggest that some combinations of basic emotions are very common in the experiments. Obviously, not all combinations are meaningful for humans. Some emotion pairs such as *happyness&anger*, *happiness&disgust*, *anger&surprise*, rarely appear together in our observations.

In summary, we observed that compound emotions are very common in the user reports, and some of the emotion pairs are highly correlated in the collected dataset. Such information will be exploited to design a machine learning algorithm for emotion detection.

5.2 Feature extraction

With the data collected from the mobile users, we need to process them into feature presentation. Specifically, the raw data is processed to different types of features including the situation of *environment*, *contact*, *APP usage*, and *activity*, which are discussed below.

5.2.1 Environment. Environmental situations are known to influence the feeling of users. The environment features can be represented by the microphone and the light sensor information from the environment. The microphone logs the sounds "heard" by the smartphone, and we take the mean and variance to indicate the volumes and dispalcements. According to [40], a sound is considered to be noise when the volume exceeds some threshold: 85-90 dBA. We further define the *noise ratio* $NR = \frac{number\ of\ noise\ samples}{total\ number\ of\ audio\ samples}$, *silence ratio* $SR = 1 - NR$ and *noise-silence ratio* $NSR = \frac{NR}{SR}$ to represent the noise condition of the environment.

Similarly, we take the mean and variance of light sensors as features, and use the *dark ratio (DR)*, *bright ratio (BR)*, *dark-bright ratio (DBR)* to represent the illumination and to infer the indoor/outdoor duration of an individual.

In terms of location information, GPS can be used to infer the location of a user outdoor. However, it cannot be used for indoor environment. On the other hand, WiFi has been used as a means of indoor positioning [13], which is also an important location information. In our system, we record the SSIDs of WiFi access points scanned by the smartphones every 5 minutes, and count the frequency each SSID appears in the WiFi log. Then we choose the frequency of the top N occurred SSIDs of each user as the indoor location feature, which approximately indicate the user's often visited indoor locations.

5.2.2 Contact. According to the study of [2], peoples' emotion states are influenced by their friends or the socially contact persons. Contact features represent the users' social connections and behaviors. From the call/SMS logs, we extract the call duration, and call/SMS frequency to represent the contact features. Call duration measures how much time the user spent on communicating with a specific contact through

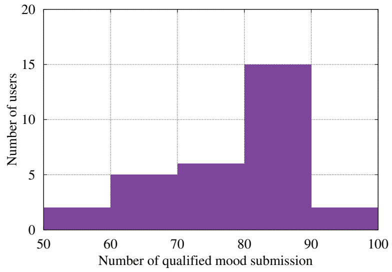


Fig. 7. Distribution of the number of qualified submissions

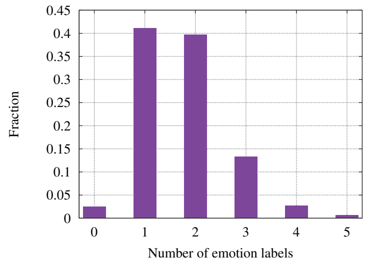


Fig. 8. Distribution of the number of emotion labels.

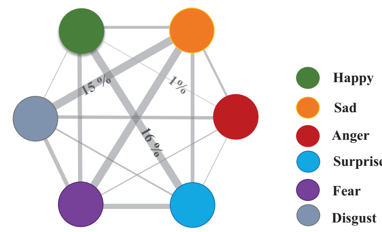


Fig. 9. Correlation of basic emotion categories.

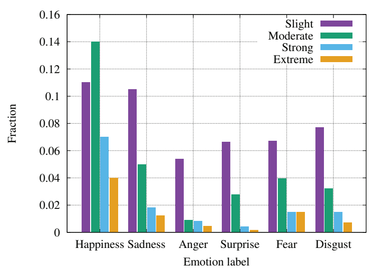


Fig. 10. Distribution of emotion levels.

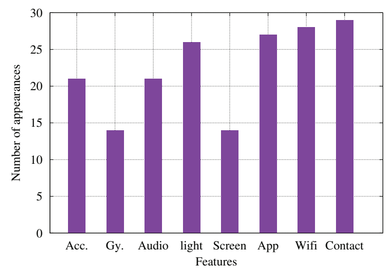


Fig. 11. Freq. of the categories of the selected features.

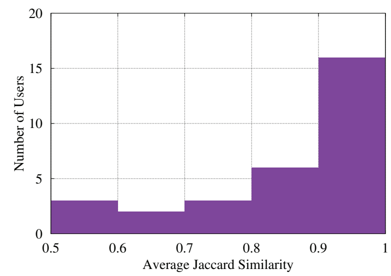


Fig. 12. Distribution of Jaccard similarity

Table 1. Description of the extracted features. The number in brackets indicates the number of the extracted features for that category. The star means that the number of extracted features is not fixed and it varies from person to person.

Type	Data source category	Extracted features description
Environment	Microphone Audio (5)	Mean, Variance, NR, SR, NSR
	Light Sensor (5)	Mean, Variance, DR, BR, DBR
	GPS (3)	GPS Longitude, Altitude, Latitude
	WiFi (20)	Frequency of SSIDs of the top 20 APs for an user
Contact	Phone Call (*)	Call frequency and duration of each contact person
	SMS (*)	SMS frequency of each contact person
APP Usage	APP log (18)	Duration of 18 APP Categories
Activity	Accelerometer (7)	Mean and Variance of three axis (X, Y, Z), Step Count
	Compass (6)	Mean and Variance of three axis (X, Y, Z)
	Gyroscope (6)	Mean and Variance of three axis (X, Y, Z)
	Screen (4)	Screen on ratio, off ratio, Sleeping Duration, Usage Amount

phonecall in a time interval. Similarly, call/SMS frequency measures how many times the user had called or sent messages to a contact during the interval.

5.2.3 APP Usage. Intuitively, people tend to use different kinds of APPs under different emotions/moods. Due to the large amount of APPs, we do not use a concrete APP as a feature. Instead, we

ALGORITHM 1: Feature selection algorithm

Input: Multi-label feature matrix M
Output: Selected feature set S
1: Transform M to k single label feature matrices
2: Use ReliefF to measure importance of features in every single label dataset
3: Output a weight vector W for all the features
4: Sum up W of all the k single label datasets
5: Sort the features according to the summed up weights
6: Choose the top N feauters as the selected feature set S

classify the APPs into 18 categories. Such category can be directly obtained from the Android Market such as Google Play[1]. The 18 categories are: *Mother and child*, *Traffic navigation*, *Image*, *Office efficiency*, *Education*, *News*, *Travel*, *Life tools*, *Life services*, *Telephone communication*, *System tools*, *Smartphone beautification*, *Social chat*, *Vedio*, *Shopping*, *Health*, *Financial management*, and *Music*. For each APP launched by the user, we use a tuple $< Category, duration >$ to denote the APP usage.

5.2.4 Activity. Activities such as going for sports or lying on the bed may imply peoples' different emotions/moods. The users' activities can somehow be derived by the accelerometer, electronic compass, and gyroscope sensors that record the human movement condition when carrying a smartphone. The accelerometer records user movement in three dimensions: X (the direction of front and back), Y (the direction of left and right), and Z (the direction of up and down) . The electronic compass records the orientation in the form of an angle with respect to magnetic north, and the gyroscope sensor record the position of the smartphone. All of them can reflect the movement states of the mobile users.

Since the accelerometer data has clear periodic patterns when the user is walking and running, such property has been used by many works to count the steps of a user [29]. In our paper, we also take *step count* as a feature of user activities.

Exploring the on/off patterns of cellphone screen leads to several interesting features, such as the *sleeping duration*, which can be estimated by the longest off-screen interval, and the phone *usage amount*, defined by the proportion of screen on-to-off duration, which indicates the time that a user spent on playing with the smartphone.

The overall features extracted from the dataset are summarized in Table 1. The integers in the brackets indicate the number of the extracted features from the type of data. Note that we do not put a specific number on the *Phone call* and *SMS* data since the number of contacts of different users are variant.

5.3 Feature selection

After feature extraction, we get 100 more features from the raw data. However, not all features play an equal role in compound emotion detection. Some features may show high correlation tp the labels; some may be less relevant; some of them may be redundant; and some may be noisy. Therefore we apply the feature selection technique to reduce the number of features to improve the efficiency of the machine learning model.

Feature selection is an important data pre-processing step in machine learning. It aims to find a subset of features $X^* \subseteq X$ to describe the dataset as well as X does, where X is the original feature space. When learning from the high-dimensional data, feature selection provides a good way to reduce the dimensions.

[1]play.google.com

Table 2. Top 6 selected features of 4 randomly users.

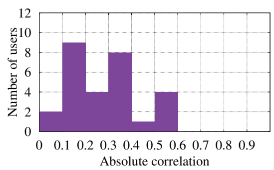
User	First Sel. Fea.	Sec. Sel. Fea.	Third Sel. Fea.	Fourth Sel. Fea.	Fifth Sel. Fea.	Sixth Sel. Fea.
1	Bright Ratio (BR)	Dark Ratio (DR)	Mean of Light Sensor Value	Variance of Audio Volume	Freq. of SSID #1873	Call Duration of Contact #1436
2	Step Count	Mean of Gyro. Z axis	Mean of Gyro. X axis	Variance of Gyro. Y axis	Duration of Shopping APP	Call Freq. of Contact #1037
3	Screen On Ratio	Screen Off Ratio	Call Freq. of Contact #1686	Call Durations of Contact #1686	Freq. of SSID #1074	Freq. of one SSID #2244
4	Freq. of SSID #3541	Freq. of SSID #5541	Freq. of SSID #1712	Duration of Social Chat APP	Duration of Video APP	Duration of Music APP

Since the reported compound emotions consists multiple basic emotions, we take each basic emotion as a label, and the compound emotion dataset can be viewed as a multi-label dataset. We adopt the *problem transformation approach* [38] for feature selection, which works as follows. Firstly, we use the *Binary Relevance* (BR) [46] method to transform the multi-label dataset to k single label datasets, where k is the number of labels (which equals 6 in our system). Secondly, we exploit *ReliefF* (RF) [39] as feature evaluation measure for each single label dataset. ReliefF outputs a weight w for each feature to represent its significance. The larger weight w is, the more important the feature is. Thirdly, for each feature we sum up its RF weights in all single label datasets, which represents its overall importance. Finally, we sort the features according to their total RF weights, and select the top N features to be used by the learning model. The details are shown in Algorithm 1.

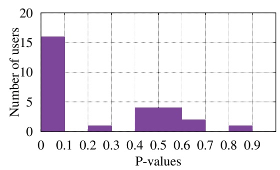
In Appendix A, we show the RF weights of all the extracted feature of a user. It shows that there are 110 features extracted from the algorithm. The weights of different features are quite different. Only very few features have weights larger than 1, while a large number of features have small weights. It suggests that a few number of features could be sufficient to build a learning model. In this paper, the default number of features used for the learning model is 6, and the influence of the number of selected features on the system performance is discussed in Section 7.4.

Table 2 shows the chosen features for 4 random users. It is interesting to see that different users have different types of features to indicate their emotional states. Three of the users' emotion states are correlated to their contacts' calls, which implies that social connections are influential to human emotions. Some of them are correlated to different APPs they used for shopping, entertainment, and social communications. User 1's emotion states are sensitive to the ambient light and sounds. The emotion states of user 2 are more related to his motion such as step count and accelerometer. User 3's emotion states seem to be more correlated with smartphone usage duration and location. While the emotion tates of user 4 are heavily related to APP usages and indoor locations. Based on the observations, different users have different set of features that correlated to their compound emotions. This suggests that personalized model should be built individually to infer the users' compound emotions. The selected features of all the 30 participants are listed in Appendix B.

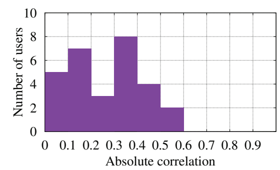
Fig. 11 shows the frequency of the categories of the selected features in the collected dataset. As shown in the result, the contact information (the social connection), the WiFi information (the indoor location information) and the APP usage are the top three most significant data sources related to users' emotion labels. This indicates that *who the user contacts, where the user has been, and which APPs the user utilizes daily* have great influence on the user's emotion states. Other important features include the light sensor (the ambient brightness), the accelerometer (individuals' physical movement status), and the



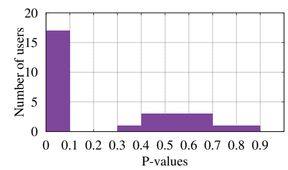
(a) Absolute correlation of the first selected feature.



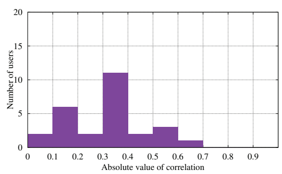
(b) P-value of the first selected feature.



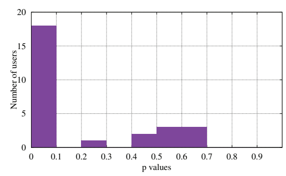
(c) Absolute correlation of the second selected feature.



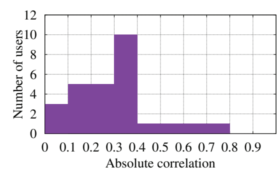
(d) P-value of the second selected feature.



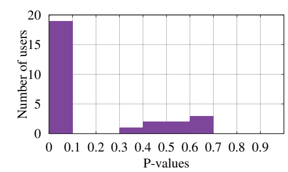
(e) Absolute correlation of the third selected feature.



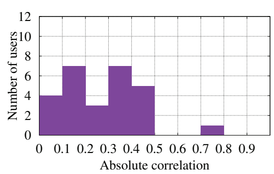
(f) P-value of the third selected feature.



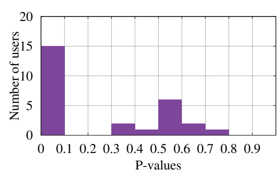
(g) Absolute correlation of the fourth selected feature.



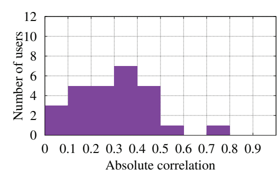
(h) P-value of the fourth selected feature.



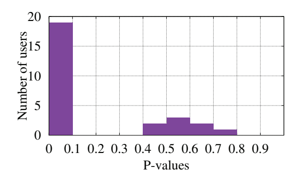
(i) Absolute correlation of the fifth selected feature.



(j) P-value of the fifth selected feature.



(k) Absolute correlation of the sixth selected feature.



(l) P-value of the sixth selected feature.

Fig. 13. Histograms of the correlation and the p-values for top 6 selected features (the number of the selected features is 6 in default, the influence of the number of selected features is discussed in section 7.4).

audio (the environmental noise), which also reflect different aspects of human activities and behaviors that have an effect on users' compound emotions.

To test whether the selected features change over time, we make observations on the individuals' feature sets selected from different time periods. Specifically, for each user, we apply the feature selection to select the top-six features in the first 20 days to form a feature set $\mathcal{F}_1$, and then we use the same method to select a feature set $\mathcal{F}_2$ in 30 days' long. We adopt the *Jaccard Coefficient* [27] to measure the similarity of the two feature set $\mathcal{F}_1$ and $\mathcal{F}_2$, which is defined as $J(\mathcal{F}_1, \mathcal{F}_2) = \frac{|\mathcal{F}_1 \cap \mathcal{F}_2|}{|\mathcal{F}_1 \cup \mathcal{F}_2|}$. The higher $J(\mathcal{F}_1, \mathcal{F}_2)$ value means the more similar between $\mathcal{F}_1$ and $\mathcal{F}_2$ and the less varying of feature sets. We show the distribution of Jaccard similarities of the 30 users in Fig. 12. As shown in the figure, half of the users have Jaccard similarity larger than 0.9, which means their feature sets are almost unchanged in a month. About 6 users' Jaccard similarities are between 0.8 and 0.9, and very few users' Jaccard similarities are less than 0.7, which implies that the feature sets of a few users change slowly over time. In summary, all the users' Jaccard similarities are larger than 0.5, which means the chosen top-six features are stable and they will not change over time dramatically. It implies that our learning model can be applied for a relative long time and it does not need to be updated frequently.

5.4 Correlation analysis

After selecting features based on Alg. 1, we conduct correlation analysis between the selected features and the emotion labels for each user. We compute the Pearson correlation coefficient [47] using the similar methods mentioned in [7][18]. We also compute the p-value associated to each correlation value. The p-value comes from the hypothesis test, in which the null hypothesis is that the correlation value equals 0. The p-value represents the probability that the null hypothesis is true. A small p-value (such as 0.1) indicates that the null hypothesis can be rejected and the correlation value is significant.

Since the emotion labels are multiple in our scenario, we calculate the correlation value between each selected feature and each emotion label respectively. A selected feature can be considered significant if it has significant correlation with at least one emotion label (i.e., $p < 0.1$). The histograms of the correlation coefficient and the p-values for the users' top six selected features are shown in Fig. 13. As shown in the result, the distributions of the absolute correlation values are uneven, which is dynamic variant between 0 to 0.6. However, one thing in common is that the p-values of the selected features are mostly concentrated on $p < 0.1$. For example, for the first selected feature shown in Fig. 13(b), about 57% users' p-values are less than 0.1. Similarly, for the second to the sixth selected feature, there are majority of users with p-values in the range (0, 0.1). Since $p < 0.1$ means significant correlation according to the theory of significant test, it confirms that the selected top six features are significantly correlated to the users' emotion states, which is reasonable to be used for compound emotion detection.

6 FACTOR GRAPH MODEL FOR COMPOUND EMOTION DETECTION

6.1 Intuition

According to the observations in section 5.1, compound emotions are the combination of the six basic emotion categories, and there exists high correlation among some pair of basic emotions such as *happiness&surprise*. A machine learning algorithm should takes into account such correlation for compound emotion detection. In this paper, we adopt the factor graph model [24] to describe the correlation between basic emotions and the correlation between features and emotion labels. Based on the proposed factor graph model, we can formulate the conditional probability of the compound emotion vector given the observed feature vector, and apply the Maximum a Posteriori (MAP) [33] principle to derive the most probable emotion labels for compound emotion. The details are introduced in the following subsections.

6.2 Model description

In the factor graph model, each node in the graph represents a variable, and the edges represent the correlation between variables, which is called factor function. Fig. 14 shows the factor graph used for compound emotion detection, which describes variety of correlations among features and emotion labels. Specially, a vector X^t is used to denote the input feature vector in the *t-th* time interval of the user's data trace, which contains N feature attributes after feature selection. And a vector $Y^t = (y_1^t, ..., y_K^t)$ is used to denote the compound emotion vector at the *t-th* time interval, where y_i^t is the *ith* emotion label and K is the number of basic emotion catalogues (K=6 for the Ekman's model). To describe the correlations between features and emotion labels, similar to the traditional multi-label learning [28], we adopt the *problem transformation approach* [38] to replicate the feature vector X^t by k copies, denoted as $\{X_k^t\}_{k=1}^K$, such that we associate each copy X_k^t with each emotion label y_k^t. With such expansion, the factor graph model of our problem is shown in Fig. 14, where each node in the upper layer corresponds to an emotion label; and each node in the lower layer corresponds to a copy of the feature vector.

The goal of the proposed factor graph model is to describe the conditional probability of users' emotional states given the features of the smartphone usage trace. Particularly, the factor graph factorizes the

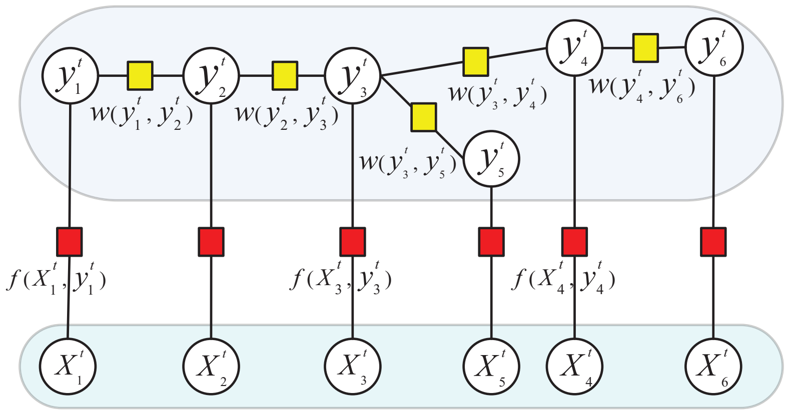


Fig. 14. The factor graph model.

global probability distribution as a product of local factor functions, each of which depends on a subset of the variables (nodes) [24][48]. Therefore, based on the observation in section 5.1, we introduce two kinds of factor functions (corresponding to the edges in the factor graph) to account for the correlations between features and emotion labels as well as the correlations across different emotion labels:

- **Feature-Label Factor Function**: $f(X_j^t, y_j^t)$ represents the correlation between the feature vector X_j^t and the emotion label y_j^t. The red rectangles in Fig. 14 represent the feature-label factor function.
- **Label-Label Factor Function**: $w(y_i^t, y_j^t)$ represents the correlation between the emotion label y_i^t and the emotion label y_j^t. The yellow rectangles in Fig. 14 represent the label-label factor function.

The factor functions can be defined in many ways to reflect the correlation factors. In this paper, we follow the fundamental Hammersley-Clifford theorem [20], which models the correlation factors by the exponential-linear functions in a Markov random field. We elaborate the factor functions in details as follows.

1) The feature-label correlation factor function. Following the Hammersley-Clifford theorem [20], we define the feature-label factor correlation function as an exponential-linear function:

$$f(X_k^t, y_k^t) = \frac{1}{Z_1} \exp\{\sum_{n=1}^{N} \alpha_{nk} \Phi(x_{nk}^t, y_k^t)\}. \tag{3}$$

The notations of the equation is explained as follows. The notation x_{nk}^t $(1 \leq n \leq N)$ indicates the n-th feature of the feature vector X_k^t. The function $\Phi(x_{nk}^t, y_k^t)$ is a binary indicator. For example, $\Phi(x_{nk}^t = \text{"}true\text{"}, y_k^t = \text{"}strong\text{"})$ means that if the n-th attribute in the feature vector (represented by x_{nk}^t) is "true" and the level of the user's k-th component in the emotion vector (represented by y_k^t) is "strong", then the indicator function value is 1, otherwise 0. The weight α_{nk} describes how strong the correlation between x_{nk}^t and y_k^t is, which is the model parameter to be learned by the machine learning algorithm. Z_1 is a normalization term to ensure the sum of the probabilities equals to 1 [44].

2) The label-label correlation factor function. Since there are multiple emotion labels, if we take into account all the possible correlations among any two labels, it will result in the explosive growth of the learning parameters, hindering the performance of our model. Moreover, accounting all emotion label pairs

ALGORITHM 2: Construction of Label Correlation Tree

Input: Label matirx M, Number of labels K
Output: Label correlation tree $G = (V, E)$
Count the times t_{ij} that label pair $< l_i, l_j >$ coexist in each row of M,
  $T = \{< l_i, l_j >, t_{ij}\}_{i,j=1\&i\neq j}^{K}$
Sort T according to the times t_{ij}
for each label pair $< l_i, l_j >$ in T **do**
  if $|E| < K - 1$ **then**
    if $l_i \in V$ & $l_j \in V$ **then**
      if $|E| = |V| - 2$ **then**
        $V = V \cup \{l_i, l_j\}$;
        $E = E \cup \{(l_i, l_j)\}$;
      else
        continue;
      end if
    else
      $V = V \cup \{l_i, l_j\}$;
      $E = E \cup \{(l_i, l_j)\}$;
    end if
  else
    break;
  end if
end for

may cause a cycling structures in the factor graph, which makes exact inference difficult [44]. Therefore, we design an algorithm to elicit the strong correlation label pairs and meanwhile avoid the cyclic structure in the factor graph.

The proposed algorithm work as follows. First, we sort the times that two emotion labels coexist in the same interval in a descending order. Then, we select the emotion label pairs one by one in accordance with the order. Each label pair corresponds to an edge in the factor graph. If an edge does not cause cyclic structure, it will be added into the factor graph; otherwise it will be discarded. The process is repeated until a tree structure called *label correlation tree* is formed. An example of the constructed label correlation tree is shown in the upper layer of the factor graph in Fig. 14. The detail algorithm is illustrated in Algorithm 2.

With the obtained label correlation tree, and following the Hammersley-Clifford theorem [20], we define the label-label correlation factor function $w(y_i^t, y_j^t)$ for each edge in the tree as:

$$w(y_i^t, y_j^t) = \frac{1}{Z_2} \exp\{\beta_{ij} \Phi(y_i^t, y_j^t)\}, \tag{4}$$

where y_i^t and y_j^t are emotion labels; $\Phi(y_i^t, y_j^t)$ is a binary indicator to show whether two emotion labels are correlated in the factor graph (1 means correlated and 0 otherwise); the weight β_{ij} quantifies the influence degree between two different emotion labels, which is the model parameter to be learned by the machine learning algorithm; and Z_2 is a normalization term to ensure the sum of the probabilities equals to 1.

ALGORITHM 3: Learning algorithm

Input: Learning rate η
Output: Model parameters θ
1: Initialize $\theta \leftarrow 0$
2: **repeat**
3: Calculate $E_D[\Phi(x_{nk}^t, y_k^t)]$ and $E_D[\Phi(y_i^t, y_j^t)]$ using the training dataset
4: Calculate $E_\theta[\Phi(x_{nk}^t, y_k^t)]$ and $E_\theta[\Phi(y_i^t, y_j^t)]$ using BP algorithm
5: Calculate the gradients $\frac{\partial L(\theta)}{\partial \theta}$ according to Eqs.(7) and (8)
6: Update parameter θ as $\theta = \theta + \eta \frac{\partial L(\theta)}{\partial \theta}$
7: **until** Convergence

6.3 Objective function

Assume there are T total time intervals in the experiment. Let $X = (X^1, ..., X^T)$ be the sequence of observed feature vectors over T time intervals, and $Y = (Y^1, ..., Y^T)$ be the sequence of the users' compound emotion vectors accordingly.

Following the common principles of factor graph [23][25][44], given the observations X and the correlation factor functions, the conditional probability of Y can be described by the product of all factor functions in the factor graph, which is expressed as

$$P(Y|X,\theta) = \prod_{t=1}^{T} \prod_{k=1}^{K} \prod_{\forall y_j^t \in \Delta(y_k^t)} f(y_k^t, X_k^t) w(y_k^t, y_j^t), \tag{5}$$

where $\Delta(y_k^t)$ is the set of emotion label nodes having an edge with y_k^t in the factor graph; and $\theta = (\{\alpha_{nk}\}, \{\beta_{ij}\})$ are model parameters to be learned.

The conditional probability given in Eq. 5 forms the objective function to be maximized, which is discussed in the next subsection.

6.4 Model learning

Given Eq. 5, we would like to determine the optimal system parameters $\theta = (\{\alpha_{nk}\}, \{\beta_{ij}\})$ to maximize the objective function (which corresponds to find an optimal mapping from the feature vector to the compound emotion vector with maximum probability). To achieve this task, the maximization problem can be transformed to minimize the following negative log-likelihood function:

$$L(\theta) = -\log P(Y|X,\theta) + \frac{\lambda}{2}(\sum_{n=1}^{N}\sum_{k=1}^{K} {\alpha_{nk}}^2 + \sum_{i=1}^{K}\sum_{j=1}^{K} {\beta_{ij}}^2), \tag{6}$$

where the last term is a L_2-regularization penalty, which is commonly used in data mining to prevent overfitting [23]; and λ is the weight of penalty factor.

We apply the gradient decent method [23] to learn the parameters θ. We obtain the gradients for our factor graph model as follows:

$$\frac{\partial L(\theta)}{\partial \alpha_{nk}} = E_\theta[\Phi(x_{nk}^t, y_k^t)] - E_D[\Phi(x_{nk}^t, y_k^t)] + \lambda \alpha_{nk}, \tag{7}$$

and

$$\frac{\partial L(\theta)}{\partial \beta_{ij}} = E_\theta[\Phi(y_i^t, y_j^t)] - E_D[\Phi(y_i^t, y_j^t)] + \lambda \beta_{ij}. \tag{8}$$

Where $E_\theta[\cdot]$ is the expectation of feature values with respect to the model parameters θ, and $E_D[\cdot]$ is the average value by counting the given pattern over the given training dataset. The learning algorithm is summarized in Algorithm 3, wherein we apply the belief propagation (BP) algorithm [15] to infer the expectation value $E_\theta[\cdot]$.

Given the learned parameter θ, we can infer user's compound emotions based on the principle of Maximum a Posteriori (MAP), i.e., finding the compound emotion vectors that maximize the likelihood given the learned parameters θ as

$$\arg\max_{y} \; P(Y = y|X, \theta). \tag{9}$$

Similarly, we use the belief propagation (BP) algorithm to calculate the marginal probabilities. Hence, the emotion vector with the highest probabilities will be obtained as the output of compound emotion detection.

7 PERFORMANCE EVALUATION

In this section, we conduct experiments on the collected dataset to analyze the performance of the proposed factor graph based compound emotion detection method.

7.1 Experiment setup

The experiments are based on the smartphone data collected from 30 students. Since different users may have different sets of features, we run the feature selection algorithm and construct the factor graph model for each user individually. By default, we use 70% data as training set to build the model, and use the remaining 30% data as the test set for performance evaluation.

We compare the performance of the proposed algorithm with three baseline classification algorithms: *Decision Tree* (DT), *Support Vector Machine* (SVM), and *Logistic Regression* (LR). We adopt the problem transformation approach that transforms the compound emotion detection problem into the detection of each dimension of basic emotion independently. To do that, we transform the obtained multi-label dataset to k single-label datasets, apply the classification algorithms on each single-label dataset separately, and combine the results to form the compound emotions. We do not utilize the conventional multi-label classifier such as ML-kNN and ML-DT due to the fact that forming a large label space for multi-label classification will lead to the short of input instances for model training.

7.2 Performance metrics

Let m be the size of the test set, and k be the number of basic emotion catalogues (k=6 in the Ekman's model). Denoted by $R = \{R_1, ..., R_m\}$ be the set of emotion reports in the test set (ground truth), where $R_i = (r_{i1}, ..., r_{ik})$ $(1 \leq i \leq m)$ be the emotion vector of the *i-th* instance in the set. Denoted by $S = \{S_1, ..., S_m\}$ be the set of emotion vectors derived from our compound emotion detection algorithm, where $S_i = (s_{i1}, ..., s_{ik})$ $(1 \leq i \leq m)$ be the *i-th* inferred emotion vector. We use the following metrics for performance evaluation in the paper.

- **Exact match:** It evaluates the percentage of emotion labels that are correctly detected by the algorithm, which is calculated by

$$ExactMatch = \frac{1}{mk} \sum_{i=1}^{m} \sum_{j=1}^{k} I(r_{ij}, s_{ij}), \tag{10}$$

where $I(\cdot)$ is an indicator function, $I(r_{ij}, s_{ij}) = 1$ when $r_{ij} = s_{ij}$, and 0 otherwise.

Table 3. Mean value of different metrics.

Mean value	DT	SVM	LR	Factor Graph
Exact match(%)	71.9	71.4	72.3	**76.0**
MAE	0.380	0.395	0.380	**0.322**
Accuracy(%)	54.4	55.9	54.7	**62.9**
Precision(%)	68.0	65.5	67.9	**80.2**
Recall(%)	60.0	64.2	60.2	**66.2**
F1-score(%)	64.1	65.8	64.1	**72.5**

- **Mean Absolute Error (MAE):** MAE is defined as the average difference between the inferred emotion vector and the ground truth, which is given by:

$$MAE = \frac{1}{m}\sum_{i=1}^{m} M(R_i, S_i), \tag{11}$$

  where $M(R_i, S_i) = \sum_{j=1}^{k} |r_{ij} - s_{ij}|$ is the absolute error between two vectors.

Like traditional multi-label learning classification, we also adopt four widely used performance metrics [50] including accuracy, precision, recall and F1-score.

- **Accuracy:** It evaluates the proportion of correctly inferred emotion labels to the total number of labels for each instance. The overall correctness is the average on all the instances, which is given by

$$Accuracy = \frac{1}{m}\sum_{i=1}^{m} \frac{|S_i \cap R_i|}{|S_i \cup R_i|}, \tag{12}$$

  where $S_i \cap R_i$ is the set of non-null common emotion labels of S_i and R_i, while $S_i \cup R_i$ is the set of non-null distinct emotion labels respectively.
- **Precision:** It is the the proportion of the correctly inferred emotion labels to the total number of inferred emotion labels for each instance. The overall precision is the average on all the instances:

$$Precision = \frac{1}{m}\sum_{i=1}^{m} \frac{|S_i \cap R_i|}{|S_i|}. \tag{13}$$

- **Recall:** It is the the proportion of the correctly inferred emotion labels to the number of actual emotion labels for one instance. The overall recall is the average on all the instances:

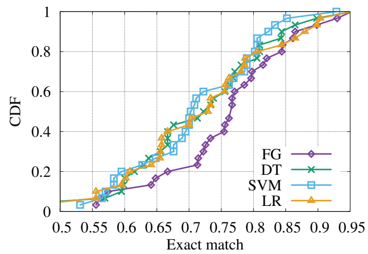
$$Recall = \frac{1}{m}\sum_{i=1}^{m} \frac{|S_i \cap R_i|}{|R_i|}. \tag{14}$$

- **F1-score:** It is a weighted harmonic mean between the precision and the recall:

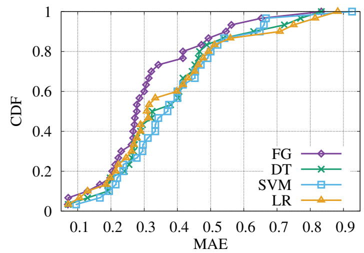
$$F1 = \frac{2 * \mathrm{Re}call * \Pr ecision}{\mathrm{Re}call + \Pr ecision}. \tag{15}$$

7.3 Numerical results

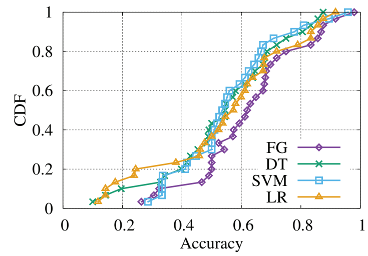
We run the factor graph based compound emotion detection algorithm and the baseline algorithms for each user in the dataset, and compute their performance metrics. The mean values of the performance metrics are compared in table 3. It is shown that the proposed algorithm outperforms the baseline algorithms in all performance metrics. For instance, the proposed compound emotion detection method



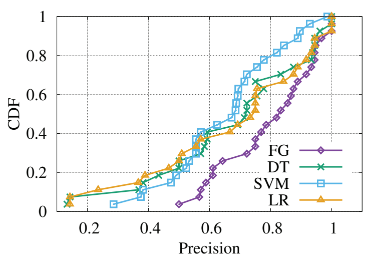
(a) CDF of Exact match.



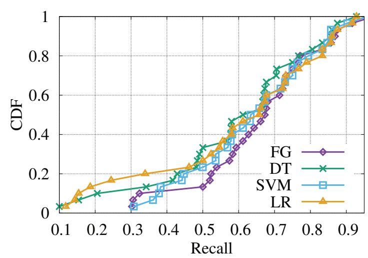
(b) CDF of MAE.



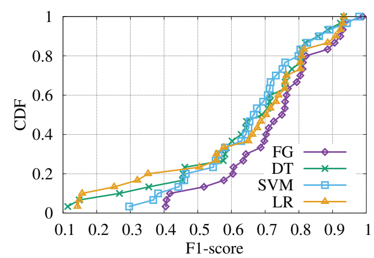
(c) CDF of Accuracy.



(d) CDF of Precision.



(e) CDF of Recall.



(f) CDF of F1-score.

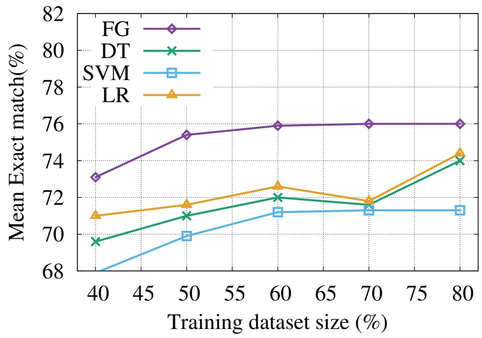
Fig. 15. CDF of metrics of different methods.

achieves 76.0% exact match on average, which has +4.1%, +4.6% and +3.7% improvement compared with DT, SVM and LR respectively. The proposed algorithm has MAE 0.322, which is much lower than that of the other algorithms (above 0.380). The accuracy, precision, recall, and F1-score of the proposed algorithm all show significant improvement compared with the baselines. The reason lies in that the propose factor graph model takes into account the correlation between emotion labels, while the baseline algorithms predict the emotion labels individually without considering their correlations.

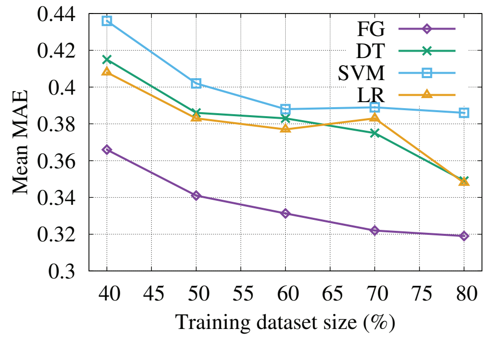
The CDF (Cumulative Distribution Function) of exact match are compared in Fig. 15(a). As shown in the figure, for the factor graph algorithm, over 80% users achieve exact match larger than 70%; but for DT, SVM, and LR models, only about 56% users have exact match larger than 70%.

Fig. 15(b) compares the CDF of MAE for different approaches in compound emotion detection. As shown in the result, for the factor graph algorithm, more than 80% users have MAE smaller than 0.4, which is much better than DT, SVM and LR. The average MAE of different approaches is shown in Table 3. The proposed factor graph method has the lowest MAE, which achieves 15%, 18% and 15% improvement compared with DT, SVM, and LR respectively.

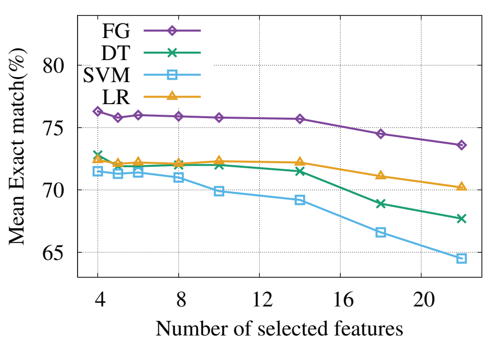
Accuracy, precision, recall and F1-score are four commonly used performance metrics in traditional muti-label learning, and their CDFs are shown in Fig. 15(c), 15(d), 15(e), 15(f) respectively. As shown in the figures, the distribution curves of the factor graph algorithm are more concentrated to the right part of the figures, which means that more users achieve higher accuracy, higher precision, higher recall, and higher F1-score compared with the baselines. The average values of the performance metrics are shown in Table 3. Compared with DT, SVM, and LR, the proposed factor graph algorithm gains +8.5%, +7.0%



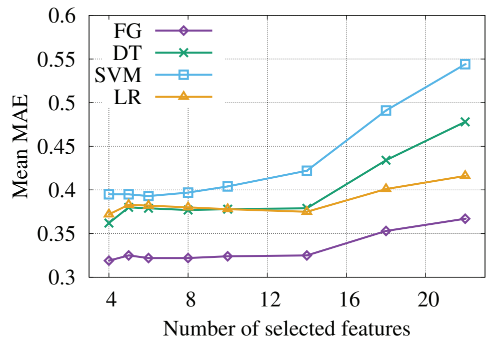
(a) Mean exact match under different training dataset size.



(b) Mean MAE under different training dataset size.



(c) Mean exact match under different number of selected features.



(d) Mean MAE under different number of selected features.

Fig. 16. Mean exact match and MAE under different training dataset size and number of selected features.

and +8.2% in average accuracy; gains +12.2%, +14.7% and +12.3% in average precision; gains +6.2%, +2.0% and +6.0% in recall; and gains +8.4%, +6.7% and +8.4% in F1-score respectively.

7.4 Parameter analysis

It should be aware that the performance of the proposed training algorithm should be influenced by several system parameters such as the chosen of training set and features. In this section, we analyze the influence of training dataset size and the number of selected features on compound emotion detection performance.

The average exact match and MAE under different training dataset size is illustrated in Fig. 16(a) and Fig. 16(b). As shown in the result, when the size of the training dataset size increases from 40% to 80%, the exact match of proposed factor graph model has +2.3% improvement and the MAE has 7% improvement. This is due to the fact that when the size of training dataset size is too small, there are not enough instances for training, which leads to underfitting. When the training dataset size is large enough, the exact match tends to be stable and could not be improved further by increasing the size of training set.

To explore the influence of the selected features, we fixed the training set size to be 70% (about 40+ instances over the all collected instances), and run the algorithm by varying the number of selected features from top-4 to top-22. The average exact match and MAE under different number of selected features is illustrated in Fig. 16(c) and Fig. 16(d). As shown in the figure, when the number of features increases, the performance decreases. When the number of selected features continues to becomes larger (¿12), there is a significant downward trend about the exact match and upward trend about the MAE. The reason is explained as follows. When the feature space becomes larger, it normally requires more training data to "feed" the model. In our experiments, the number of instance is about 40+, which is relative small. Increasing the number of features will easily cause over-fitting of the model, which will decrease the performance. According to the figure, the performance is almost the same when the number of features from 4 to 6, which implies that a small number of significant features are enough for the detection task. Note that too few features will also cause under-fitting or lack of generalization in machine learning theories. We choose 6 as the default number of features in our experiments since it performs well under different conditions and can avoid over-fitting and under-fitting issues.

In summary, the proposed compound emotion detection method yields the highest exact match and the lowest MAE compared with the baseline approaches.

8 CONCLUSION

Automatic emotion detection is important to build intelligent systems and to evaluate individual's mental well-being. In this paper, we propose MoodExplorer, a system to detect people's compound emotion based on smartphone sensing data. We first present the compound emotion model that models people's compound emotion by the combination of basic emotion categories, and formulate the compound emotion detection problem as a multi-label classification problem. Then we conduct feature extraction, feature selection, and correlation analysis on the user reported dataset, which shows that compound emotions frequently appear and have high correlation with some smartphone usage features such as the sensor data and APP usage patterns. We further introduce a factor graph model to describe the correlations among emotion labels and features, and propose a machine learning algorithm for compound emotion detection. The proposed compound emotion detection mechanism was implemented in an Android APP called MoodExplorer. The APP ran on 30 university students for one month, and their smartphone sensing data was collected for analysis. We conduct extensive experiments on the collected dataset, which show that MoodExplorer can infer user's compound emotion with exact match of 76.0% on average.

The study of smartphone-based compound emotion detection in our paper achieves several cheerful results, which include an interesting research problem of compound emotion detection, a well-developed machine learning model, and a decent accuracy for compound emotion detection. However, we shall also mention some limits of the work and the possible future directions. First, smartphone-based sensing is unobtrusive but intermittent and coarse-grained. Since the users may not carry the smartphones all day long, the sensing data only reflects parts of the human activities. Therefore it can roughly estimate user's compound emotion in a coarse time interval. Towards fine-grained short-term emotion detection is still a challenging task. Second, the proposed model is only tested on a small scale dataset. The result is verified by 30 university students, and it is unclear whether it can be generalized to people under different occupancy, different age, different culture, and different races. The experiment could be biased, and the collected data and emotion reports could be distortion since we could not tell whether a subject reports his/her emotions truthfully and thoroughly. Third, the proposed detection model is person-specific and need to be trained individually. It requires a relative long period to collect data from the user to build the model, which prevents it from being widely deployed in practice. In the next step, we may seek a general model for user groups by clustering the features and users, and we may apply the semi-supervised learning and transfer learning techniques to reduce the startup time.

The research on emotion detection based on smartphone sensing is a very young area, and it enables a new way to sense, understand, interact, and intervene on human emotional feeling and mental wellbeing. It needs the interdisciplinary efforts from sociology, psychology, neuroscience, and computer science. Together they will make more progress towards automatic emotion recognition and prediction.

ACKNOWLEDGEMENT

The authors sincerely acknowledge the reviewers and editors who provided valuable comments and feedbacks to help to improve the quality of the paper. This work was partially supported by the National Key R&D Program of China (Grant No. 2017YFB1001800), the National Natural Science Foundation of China (Grant Nos. 61672278, 61373128, 61321491), the Collaborative Innovation Center of Novel Software Technology and Industrialization, and the Sino-German Institutes of Social Computing.

REFERENCES

[1] J. Ang, R. Dhillon, A. Krupski, E. Shriberg, and A. Stolcke. Prosody-based automatic detection of annoyance and frustration in human-computer dialog. In *INTERSPEECH*. Citeseer, 2002.
[2] S. Aral and D. Walker. Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341, 2012.
[3] A. B. Ashraf, S. Lucey, J. F. Cohn, T. Chen, Z. Ambadar, K. M. Prkachin, and P. E. Solomon. The painful face–pain expression recognition using active appearance models. *Image and vision computing*, 27(12):1788–1796, 2009.
[4] X. Bao, S. Fan, A. Varshavsky, K. Li, and R. Roy Choudhury. Your reactions suggest you liked the movie: Automatic content rating via reaction sensing. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 197–206. ACM, 2013.
[5] C. Beedie, P. Terry, and A. Lane. Distinctions between emotion and mood. *Cognition & Emotion*, 19(6):847–878, 2005.
[6] A. Bogomolov, B. Lepri, M. Ferron, F. Pianesi, and A. S. Pentland. Daily stress recognition from mobile phone data, weather conditions and individual traits. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 477–486. ACM, 2014.
[7] L. Canzian and M. Musolesi. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pages 1293–1304. ACM, 2015.
[8] B. Cao, L. Zheng, C. Zhang, P. S. Yu, A. Piscitello, J. Zulueta, O. Ajilore, K. Ryan, and A. D. Leow. Deepmood: Modeling mobile phone typing dynamics for mood detection. 2017.
[9] J. R. Crawford and J. D. Henry. The positive and negative affect schedule (panas): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43(3):245–265, 2004.
[10] E. Diener, S. Oishi, and R. E. Lucas. Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual review of psychology*, 54(1):403–425, 2003.
[11] S. Du, Y. Tao, and A. M. Martinez. Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences*, 111(15):E1454–E1462, 2014.
[12] P. Ekman, W. V. Friesen, M. O'Sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, P. E. Ricci-Bitti, et al. Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, 53(4):712, 1987.
[13] A. Exler, M. Urschel, A. Schankin, and M. Beigl. Smartphone-based detection of location changes using wifi data. In *International Conference on Wireless Mobile Communication and Healthcare*, pages 164–167. Springer, 2016.
[14] J. H. Fowler, N. A. Christakis, et al. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *British Medical Journal*, 337:a2338, 2008.
[15] B. J. Frey and D. J. MacKay. A revolution: Belief propagation in graphs with cycles. *Advances in neural information processing systems*, pages 479–485, 1998.
[16] T. Geller. How do you feel?: Your computer knows. *Communications of the ACM*, 57(1):24–26, 2014.
[17] A. Gluhak, M. Presser, L. Zhu, S. Esfandiyari, and S. Kupschick. Towards mood based mobile services and applications. In *European Conference on Smart Sensing and Context*, pages 159–174. Springer, 2007.

[18] J. Golbeck, C. Robles, and K. Turner. Predicting personality with social media. In *CHI'11 extended abstracts on human factors in computing systems*, pages 253–262. ACM, 2011.

[19] J. J. Gross and O. P. John. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology*, 85(2):348, 2003.

[20] J. M. Hammersley and P. Clifford. Markov fields on finite graphs and lattices. 1971.

[21] K. C. Herdem. Reactions: Twitter based mobile application for awareness of friends' emotions. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 796–797. ACM, 2012.

[22] J. Hernandez, M. E. Hoque, W. Drevo, and R. W. Picard. Mood meter: counting smiles in the wild. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 301–310. ACM, 2012.

[23] D. Koller and N. Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[24] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.

[25] J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the eighteenth international conference on machine learning (ICML '01)*, volume 1, pages 282–289, 2001.

[26] H. Leng, Y. Lin, and L. Zanzi. An experimental study on physiological parameters toward driver emotion recognition. In *International Conference on Ergonomics and Health Aspects of Work with Computers*, pages 237–246. Springer, 2007.

[27] M. Levandowsky and D. Winter. Distance between sets. *Nature*, 234(5323):34–35, 1971.

[28] S. Li, L. Huang, R. Wang, and G. Zhou. Sentence-level emotion classification with label and context dependence. *Proceedings of ACL-2015*, pages 1045–1053, 2013.

[29] W. Li, Y. Hu, X. Fu, S. Lu, and D. Chen. Cooperative positioning and tracking in disruption tolerant networks. *IEEE Transactions on Parallel and Distributed Systems*, 26(2):382–391, 2015.

[30] R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong. Moodscope: building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 389–402. ACM, 2013.

[31] G. MacKerron and S. Mourato. Happiness is greater in natural environments. *Global Environmental Change*, 23(5):992–1000, 2013.

[32] A. Mottelson and K. Hornbæk. An affect detection technique using mobile commodity sensors in the wild. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 781–792. ACM, 2016.

[33] K. P. Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[34] R. Plutchik. A general psychoevolutionary theory of emotion. *Theories of emotion*, 1(3-31):4, 1980.

[35] J. Posner, J. A. Russell, and B. S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(03):715–734, 2005.

[36] R. Reisenzein, M. Studtmann, and G. Horstmann. Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emotion Review*, 5(1):16–23, 2013.

[37] S. Servia-Rodríguez, K. K. Rachuri, C. Mascolo, P. J. Rentfrow, N. Lathia, and G. M. Sandstrom. Mobile sensing at the service of mental well-being: a large-scale longitudinal study. In *Proceedings of the 26th International Conference on World Wide Web*, pages 103–112. International World Wide Web Conferences Steering Committee, 2017.

[38] N. SpolaôR, E. A. Cherman, M. C. Monard, and H. D. Lee. A comparison of multi-label feature selection methods using the problem transformation approach. *Electronic Notes in Theoretical Computer Science*, 292:135–151, 2013.

[39] N. Spolaôr, E. A. Cherman, M. C. Monard, and H. D. Lee. Relieff for multi-label feature selection. In *Intelligent Systems (BRACIS), 2013 Brazilian Conference on*, pages 6–11. IEEE, 2013.

[40] S. A. Stansfeld and M. P. Matheson. Noise pollution: non-auditory effects on health. *British medical bulletin*, 68(1):243–257, 2003.

[41] T. Stütz, T. Kowar, M. Kager, M. Tiefengrabner, M. Stuppner, J. Blechert, F. H. Wilhelm, and S. Ginzinger. Smartphone based stress prediction. In *International Conference on User Modeling, Adaptation, and Personalization*, pages 240–251. Springer, 2015.

[42] Y. Suhara, Y. Xu, and A. Pentland. Deepmood: Forecasting depressed mood based on self-reported histories via recurrent neural networks. In *Proceedings of the 26th International Conference on World Wide Web*, pages 715–724. International World Wide Web Conferences Steering Committee, 2017.

[43] B. Sun, Q. Ma, S. Zhang, K. Liu, and Y. Liu. iself: Towards cold-start emotion labeling using transfer learning with smartphones. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 1203–1211. IEEE, 2015.

[44] C. Sutton and A. McCallum. An introduction to conditional random fields. *arXiv preprint arXiv:1011.4088*, 2010.

[45] S. S. Tomkins. Affect, imagery, consciousness: Vol. i. the positive affects. 1962.

[46] G. Tsoumakas, I. Katakis, and I. Vlahavas. Mining multi-label data. In *Data mining and knowledge discovery handbook*, pages 667–685. Springer, 2009.
[47] L. Wasserman. *All of statistics: a concise course in statistical inference*. Springer Science & Business Media, 2013.
[48] Y. Yang, J. Jia, B. Wu, and J. Tang. Social role-aware emotion contagion in image social networks. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
[49] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE transactions on pattern analysis and machine intelligence*, 31(1):39–58, 2009.
[50] M.-L. Zhang and Z.-H. Zhou. A review on multi-label learning algorithms. *IEEE transactions on knowledge and data engineering*, 26(8):1819–1837, 2014.
[51] S. Zhang and P. Hui. A survey on mobile affective computing. *ArXiv Prepr. ArXiv14101648*, 2014.
[52] Y. Zhang, J. Tang, J. Sun, Y. Chen, and J. Rao. Moodcast: Emotion prediction via dynamic continuous factor graph model. In *Proceedings of the 10th International Conference on Data Mining*, pages 1193–1198. IEEE, 2010.
[53] M. Zhao, F. Adib, and D. Katabi. Emotion recognition using wireless signals. In *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking (MobiCom '16)*, pages 95–108, 2016.
[54] S. Zhao, H. Yao, and X. Jiang. Predicting continuous probability distribution of image emotions in valence-arousal space. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 879–882. ACM, 2015.
[55] Y. Zhou, H. Xue, and X. Geng. Emotion distribution recognition from facial expressions. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 1247–1250. ACM, 2015.

Appendix

A THE RF WEIGHTS OF ALL FEATURES OF A USER

Fea. ID	Feature	Weight	Fea. ID	Feature	Weight
1	Dark Ratio (DR)	1.35	2	Var. of Acc. X axis	1.23
3	Var. of Acc. Y axis	1.15	4	Dur. of Travel APP	1.13
5	Dur. of Life Services APP	1.13	6	Mean of Acc. Z axis	1.10
7	Bright Ratio (BR)	1.01	8	Dark Bright Ratio (DBR)	0.99
9	Freq. of SSID #2009	0.93	10	Freq. of SSID #5455	0.93
11	Freq. of SSID #2060	0.92	12	Call Freq. of Contact #1870	0.83
13	Silence Ratio (SR)	0.75	14	Mean of Gyro. X axis	0.73
15	Var. of Audio Volume	0.64	16	Noise Ratio (NR)	0.62
17	NSR	0.58	18	Call Duration of Contact #1870	0.58
19	Call Freq. of Contact #2021	0.56	20	Var. of Acc. Z axis	0.48
21	Var. of Gyro. Z axis	0.45	22	Mean of Gyro. Y axis	0.45
23	Var. of Gyro. X axis	0.44	24	Mean of Audio Volume	0.40
25	Mean of Gyro. Z axis	0.32	26	Mean of Acc. X axis	0.30
27	Call Duration of Contact #6387	0.29	28	Call Duration of Contact #1577	0.29
29	Call Duration of Contact #2243	0.29	30	Call Freq. of Contact #1577	0.29
31	Call Duration of Contact #3367	0.28	32	Call Freq. of Contact #3367	0.28
33	Call Freq. of Contact #1713	0.28	34	Call Duration of Contact #1713	0.28
35	Call Freq. of Contact #2243	0.28	36	Call Freq. of Contact #6387	0.28
37	Call Freq. of Contact #2241	0.28	38	Call Duration of Contact #5426	0.28
39	Call Freq. of Contact #5426	0.28	40	Call Duration of Contact #2021	0.27
41	Mean of Acc. Y axis	0.27	42	Var. of Gyro. Y axis	0.18
43	Call Freq. of Contact #1905	0.17	44	Call Duration of Contact #4127	0.14
45	Call Freq. of Contact #1683	0.13	46	Var. of Light Sensor Value	0.13
47	Dur. of Mother&Child APP	0.13	48	Dur. of Smartphone Beautification APP	0.13
49	Freq. of SSID #1712	0.13	50	Freq. of SSID #2244	0.13
51	Freq. of SSID #5135	0.13	52	Freq. of SSID #1131	0.13
53	Dur. of Image APP	0.13	54	Dur. of Vedio APP	0.13
55	Dur. of Office Efficiency APP	0.11	56	Dur. of Shopping APP	0.11
57	Step Count	0.11	58	GPS Altitude	0.09
59	GPS Latitude	0.08	60	GPS Longitude	0.08
61	Dur. of System Tools APP	0.08	62	Call Duration of Contact #1683	0.08
63	Call Duration of Contact #2241	0.03	64	Dur. of Life Tools APP	0.02
65	Dur. of Financial Management APP	0.02	66	Dur. of Social Chat APP	0.00
67	Dur. of Health APP	0.00	68	SMS Freq. of Contact #6387	0.00
69	SMS Freq. of Contact #1577	0.00	70	SMS Freq. of Contact #3367	0.00
71	Freq. of SSID #2116	0.00	72	Freq. of SSID #1478	0.00
73	Freq. of SSID #9034	0.00	74	Freq. of SSID #5458	0.00
75	Freq. of SSID #3546	0.00	76	Dur. of Telephone Communication APP	0.00
77	Call Duration of Contact #1905	0.00	78	SMS Freq. of Contact #1683	0.00

Fea. ID	Feature	Weight	Fea. ID	Feature	Weight
79	SMS Freq. of Contact #1870	0.00	80	SMS Freq. of Contact #2021	0.00
81	Dur. of News APP	0.00	82	Dur. of Traffic Navigation APP	0.00
83	Dur. of Education APP	0.00	84	Dur. of Music APP	0.00
85	SMS Freq. of Contact #2243	-0.06	86	SMS Freq. of Contact #2241	-0.06
87	SMS Freq. of Contact #5426	-0.06	88	SMS Freq. of Contact #1905	-0.06
89	Mean of Magnetic Z axis	-0.09	90	Mean of Magnetic Y axis	-0.09
91	Var. of Magnetic Y axis	-0.09	92	Call Freq. of Contact #4127	-0.11
93	Mean of Light Sensor Value	-0.11	94	Freq. of SSID #1618	-0.11
95	Freq. of SSID #1627	-0.11	96	Freq. of SSID #1875	-0.11
97	Freq. of SSID #1751	-0.12	98	Var. of Magnetic X axis	-0.12
99	Mean of Magnetic X axis	-0.12	100	Var. of Magnetic Z axis	-0.12
101	SMS Freq. of Contact #1713	-0.12	102	SMS Freq. of Contact #4127	-0.13
103	Freq. of SSID #9673	-0.14	104	Freq. of SSID #6470	-0.14
105	Freq. of SSID #4707	-0.14	106	Freq. of SSID #3907	-0.16
107	Sleeping Duration	-0.16	108	Screen on Ratio	-0.16
109	Usage Amount	-0.17	110	Screen off Ratio	-0.19

Table 4. Example: the weight of an individual's all features.

B THE SELECTED FEATURES FOR ALL 30 USERS

User	First Sel. Fea.	Sec. Sel. Fea.	Third Sel. Fea.	Fourth Sel. Fea.	Fifth Sel. Fea.	Sixth Sel. Fea.
1	Bright Ratio (BR)	Dark Ratio (DR)	Mean of Light Sensor Value	Var. of Audio Volume	Freq. of SSID #1873	Call Dur. of Contact #1436
2	Step Count	Mean of Gyro. Z axis	Mean of Gyro. X axis	Variance of Gyro. Y axis	Duration of Shopping APP	Call Freq. of Contact #1037
3	Screen On Ratio	Screen Off Ratio	Call Freq. of Contact #1686	Call Dur. of Contact #1686	Freq. of SSID #1074	Freq. of one SSID #2244
4	Freq. of SSID #3541	Freq. of SSID #5541	Freq. of SSID #1712	Dur. of Social Chat APP	Duration of Video APP	Duration of Music APP
5	Call Dur. of Contact #4890	Call Freq. of Contact #1682	Call Dur. of Contact #1133	Call Freq. of Contact #1803	Call Dur. of Contact #3769	SMS Freq. of Contact #2110
6	Bright Ratio (BR)	Freq. of SSID #1074	Freq. of SSID #4392	Freq. of SSID #2244	Noise Ratio (NR)	Dark Ratio (DR)
7	Dark Bright Ratio (DBR)	Duration of Education App	Mean of Gyro. Y axis	Mean of Acc. Y axis	Dark Ratio (DR)	Screen On Ratio
8	Silence Ratio (SR)	Call Freq. of Contact #1199	Dark Ratio (DR)	SMS Freq. of Contact #1086	Mean of Audio Volume	Variance of Acc. X axis
9	Step Count	Mean of Gyro. Y axis	Mean of Acc. X axis	Variance of Audio Volume	SMS Freq. of Contact #1829	Dark Ratio (DR)
10	Freq. of SSID #2242	Freq. of SSID #1904	Freq. of SSID #7676	Dur. of System Tools APP	Screen On Ratio	Sleeping Duration
11	Silence Ratio (SR)	Mean of Gyro. Z axis	Bright Ratio (BR)	Mean of Audio Volume	SMS Freq. of Contact #1776	Freq. of SSID #1576
12	Dark Ratio (DR)	Variance of Acc. X axis	Variance of Acc. Y axis	Duration of Travel APP	Dur. of Life Services APP	Mean of Acc. Z axis
13	Freq. of SSID #6388	Mean of Light Sensor Value	Call Dur. of Contact #2112	Bright Ratio (BR)	Freq. of one SSID #2244	Variance of Audio Volume
14	Mean of Acc. Z axis	Mean of Acc. Y axis	Dark Ratio (DR)	Mean of Acc. X axis	Mean of Gyro. X axis	Freq. of SSID #1869
15	Freq. of SSID #5427	Freq. of SSID #8244	Variance of Acc. Y axis	Step Count	Usage Amount	Duration of Education APP
16	Duration of Vedio APP	Duration of Music APP	Duration of Social Chat APP	Variance of Acc. X axis	Duration of News APP	Variance of Acc. Z axis
17	Freq. of SSID #8224	Screen On Ratio	Sleeping Duration	Call Freq. of Contact #2076	SMS Freq. of Contact #2076	Call Durations of Contact #1598
18	Mean of Acc. Z axis	Freq. of one SSID #1712	SMS Freq. of Contact #3889	Variance of Acc. Z axis	Screen On Ratio	Mean of Acc. Y axis
19	Silence Ratio (SR)	Bright Ratio (BR)	Duration of Shopping APP	Dur. of Traffic Navigation APP	Duration of News APP	Noise Ratio (NR)
20	Noise Ratio (NR)	Mean of Audio Volume	Variance of Audio Volume	Freq. of SSID #2022	Screen Off Ratio	Screen On Ratio
21	Audio Mean	Freq. of SSID #1025	Noise Ratio (NR)	Duration of Education APP	Screen On Ratio	Silence Ratio (SR)

22	Freq. of SSID #3368	Mean of Gyro. X axis	Dark Ratio (DR)	Mean of Gyro. Y axis	Variance of Gyro. X axis	Dur. of Office Efficiency APP
23	Duration of M&C APP	Duration of Shopping APP	Duration of Vedio APP	Screen Off Ratio	Screen On Ratio	Mean of Gyro. Y axis
24	Variance of Acc. Y axis	Duration of Social Chat APP	Duration Financial APP	Dur. of System Tools APP	Freq. of SSID #1684	Duration of Education APP
25	Screen Off Ratio	Screen On Ratio	Freq. of SSID #5427	Mean of Gyro. Z axis	Mean of Acc. Z axis	Call Freq. of Contact #8274
26	Freq. of SSID #4126	Call Freq. of Contact #4348	Call Dur. of Contact #1118	Duration of Social Chat APP	Bright Ratio (BR)	Var. of Light Sensor Value
27	Dark Ratio (DR)	Call Freq. of Contact #1367	SMS Freq. of Contact #1367	SMS Freq. of Contact #1595	Mean of Light Sensor Value	Variance of Audio Volume
28	Noise Silence Ratio (NSR)	Bright Ratio (BR)	Mean of Acc. Z axis	Call Freq. of Contact #1897	Variance of Audio Volume	Dark Ratio (DR)
29	Freq. of SSID #2242	Freq. of SSID #6388	Mean of Gyro. X axis	Bright Ratio (BR)	Dark Ratio (DR)	Mean of Gyro. Y axis
30	Noise Ratio (NR)	Dark Ratio (DR)	Duration of System Tools APP	SMS Freq. of Contact #4122	Call Dur. of Contact #4122	Call Dur. of Contact #6463

Table 5. Selected features of another 20 users.

Received May 2017; revised August 2017; accepted October 2017